

# Prototype-Based Fuzzy Clustering melalui Algoritma FCM pada Pengklasteran Kabupaten-Kabupaten di Jawa Timur berdasarkan Karakteristik Perempuan

HADI TANUJI

Jurusan Matematika, FMIPA, Universitas Sriwijaya, Indonesia

**Intisari:** Keberadaan outlier menjadi masalah serius pada analisis kluster. Metode kluster tradisional yang sederhana seperti metode hirarkhi pada akhirnya menjadi tidak baik karena hasilnya akan berbeda-beda dan tidak memiliki ketahanan (*robust*) pada struktur kluster. Upaya paling mudah mengatasinya adalah dengan membuang data outlier tersebut, tetapi ini bukan jalan terbaik. Upaya yang dapat dilakukan adalah mencari metode alternatif untuk mengatasi adanya outlier tanpa harus membuangnya. Salah satu yang dapat digunakan adalah Prototype-Based Fuzzy Clustering melalui Algoritma Fuzzy C-Mean. Makalah ini membahas penggunaan Prototype-Based Fuzzy Clustering melalui Algoritma Fuzzy C-Mean untuk mengatasi kasus outlier. Hasilnya dibandingkan dengan metode agglomerative. Pengklasteran didasarkan pada data karakteristik perempuan hasil Survey Ekonomi Nasional tahun 2006. Berdasarkan hasil yang diperoleh dengan Prototype-Based Fuzzy Clustering, struktur kluster menjadi lebih robust terhadap data outlier.

**Kata kunci:** Prototype-Based Fuzzy Clustering, Algoritma, Fuzzy C-Mean, Agglomerative, Pencilan

**Email:** htanuji@yahoo.com

## 1 PENDAHULUAN

Salah satu masalah pada penelitian yang melibatkan banyak peubah maupun objek adalah upaya mencari analisis yang dapat menyederhanakan permasalahan yang rumit terkait dengan banyaknya peubah tersebut. Johnson dan Wichern (2007) menguraikan masalah analisis multivariate sebagai bagian dari upaya menangani model dengan jumlah variable yang besar.

Salah satu analisis multivariate berkaitan dengan upaya penyederhanaan masalah tersebut adalah analisis kluster (*cluster analysis*), yang dapat mereduksi data dengan cara mengidentifikasi sejumlah kelompok yang lebih kecil dari keseluruhan data. Tujuan identifikasi kelompok ini adalah untuk menyatukan elemen-elemen atau objek penelitian yang mempunyai kesamaan sifat, atribut atau karakteristik lebih dekat ke dalam satu kelompok. Dengan demikian analisis kluster pada prosesnya akan mengelompokkan objek (*respondents, products, firms, variables, etc.*) ke dalam beberapa kelompok atau kluster sehingga sifat-sifat dari objek-objek yang berada pada satu kelompok akan mirip dan objek-objek yang berada pada kelompok yang berbeda mempunyai sifat berbeda.

Permasalahan berikutnya muncul ketika upaya untuk melakukan pengelompokan menjadi sulit dilakukan karena dipengaruhi banyak faktor, yang

salah satunya adalah adanya data pencilan (*outlier*). Metode kluster tradisional yang sederhana seperti metode hirarkhi pada akhirnya menjadi tidak baik karena hasilnya akan berbeda-beda dan tidak memiliki ketahanan (*robust*) pada struktur kluster.

Beberapa pendekatan metodel kluster telah dilakukan untuk menghasilkan kluster yang lebih memiliki ketahanan (*robust*), diantaranya: metode klasterisasi fuzzy berbasis prototype (*prototype-based fuzzy clustering*) (Bezdek, 1981), Ohashi Algorithm (Ohashi, 1984), *noise clustering (NC) method*, *possibilistic clustering method* (Krishnapuram dan Keller, 1993), *the mountain method* (Yager dan Filev, 1994), *the deterministic annealing/least biased fuzzy clustering method* (Beni dan Liu, 1994), dan Probabilistic Neural Network-based robust clustering (Glotsos dkk, 2004).

Pada makalah ini akan ditampilkan penerapan metode klasterisasi fuzzy berbasis prototype pada data karakteristik perempuan di Provinsi Jawa Timur.

## 2 KAJIAN LITERATUR

### Analisis Kluster

Analisis kluster adalah suatu metode yang bertujuan mengelompokkan objek (*respondents, products, firms, variables, etc.*) ke dalam beberapa kelompok atau kluster sehingga sifat-sifat dari objek-objek yang

berada pada satu kelompok akan mirip dan objek-objek yang berada pada kelompok yang berbeda mempunyai sifat berbeda (Johnson dan Wicern, 2007).

Konsep dasar penggunaan analisis kluster adalah penghitungan kemiripan atau ukuran jarak antar objek. Beberapa jenis ukuran jarak yang dapat digunakan adalah jarak Euclid, jarak mahalanobis, Minscowki metric, Canberra metric dan beberapa ukuran jarak lainnya. Ukuran jarak yang biasa digunakan adalah jarak *Euclid* dan *Mahalanobis*. Jarak Euclid digunakan jika informasi mengenai sebaran data asal tidak diketahui dan peubah-peubah yang diamati tidak berkorelasi.

Misalkan matriks  $X = (X_1, X_2, \dots, X_n)$ , maka jarak Euclid kuadrat antara  $x_i$  dan  $x_j$  didefinisikan sebagai:

$$d_{ij}^2 = (x_i - x_j)^T (x_i - x_j) \quad (1)$$

Jika antara peubah yang diamati terdapat korelasi maka dapat dilakukan transformasi data asal menjadi komponen utamanya. Jika tidak dilakukan transformasi maka digunakan ukuran jarak Mahalanobis yang didefinisikan sebagai:

$$d_{ij}^2 = (x_i - x_j)^T S^{-1} (x_i - x_j) \quad (2)$$

dengan  $S^{-1}$  adalah matriks kovarians sampel.

Saat ini terdapat 2 tipe metode kluster yaitu algoritma berhirarkhi dan algoritma *partitioning*. Algoritma berhirarkhi dibagi dalam algoritma *agglomerative* (penggabungan) dan *splitting* (*divisive*) (Hardle dan Simar, 2007). Proses pengelompokan (klasterisasi) pada cara *agglomerative* dimulai dengan melihat objek secara individu-individu. Jadi dimulai dengan adanya kelompok sebanyak objek penelitian. Objek-objek yang mempunyai derajat kemiripan atau jarak paling dekat dimasukkan ke dalam kelompok yang sama terlebih dahulu. Kelompok-kelompok yang telah terbentuk pada tahap pertama ini digabungkan lagi sesuai dengan derajat kemiripan yang dipunyainya. Pada akhirnya, kluster-kluster yang mempunyai derajat kemiripan atau jarak yang sangat dekat digabungkan sehingga elemen-elemen atau objek dari kluster-kluster tersebut juga digabungkan dalam satu kluster. Cara *agglomerative* mempunyai banyak metode, antara lain *single linkage*, *average linkage*, *centroid linkage*, *complete linkage*, *median*, dan *ward linkage*.

Proses klasterisasi dengan cara *divisive* dimulai dengan membangun satu kluster besar yang berisi seluruh objek penelitian. Kluster tersebut mula-mula dipecah menjadi dua sub kluster. Isi dari satu sub kluster yang terbentuk mempunyai jarak atau deraj

at kemiripan yang paling jauh dengan isi dari sub kluster yang lain. Sub kluster-sub kluster tersebut kemudian dibagi kembali ke dalam sub kluster yang baru yang berisi objek yang mempunyai jarak yang saling berjauhan. Proses ini terus berjalan sampai terbentuk sub kluster sebanyak objek penelitian, atau sampai setiap objek menjadi sub kluster.

Sementara itu algoritma *partitioning* dimulai dari sebuah kluster yang telah didefinisikan dan diproses melalui perubahan elemen antar kluster sampai dihasilkan skor tertentu yang paling optimal. Dua teknik *partitioning* yang sering digunakan adalah *hard* atau *crisp clustering* dan *fuzzy clustering*. *Hard clustering* mengelompokkan setiap titik data kepada tepat satu kluster sedangkan *fuzzy clustering* mengelompokkan setiap titik data ke dalam beberapa kluster dengan derajat keanggotaan yang bervariasi.

Perbedaan utama teknik *partitioning* dari teknik berhirarkhi adalah bahwa penempatan objek atau elemen ke dalam kluster-kluster mungkin berubah selama aplikasi algoritma, sementara pada teknik berhirarkhi sekali sebuah kluster telah ditemukan dan elemen telah ditempatkan dalam sebuah kluster, maka penempatan elemen ini tidak berubah. (Hardle dan Simar, 2007).

### Agglomerative Hierarchical Clustering

1. Diawali dengan semua observasi dianggap sebagai cluster. Anggap banyaknya  $k=n$  objek, terdapat matriks jarak antar objek berukuran  $N \times N$ ,  $D = \{d_{ik}\}$
2. Dengan menggunakan ukuran jarak yang telah dipilih, cari dua kluster yang paling mirip (jarak paling kecil), misalnya ukuran jarak yang paling dekat adalah antara cluster U dan V, yaitu  $d_{UV}$
3. Gabungkan U dan V menjadi kluster baru UV. Berikutnya diperbarui lagi matrik jarak: a) menghapus baris dan kolom yang berhubungan dengan U dan V, b) menambahkan baris dan kolom dengan menghitung jarak UV dengan cluster yang lain.
4. Ulangi langkah 2 dan 3 total  $N-1$  kali. Sampai semua objek menjadi single cluster.

Beberapa prosedur hierarchial *agglomerative* yang dapat digunakan adalah:

- Single Linkage (nearest neighbor)  
 $D_{(UV)W} = \min\{d_{UW}, d_{VW}\}$
- Complete Linkage (farthest neighbor)  
 $D_{(UV)W} = \max\{d_{UW}, d_{VW}\}$
- Average Linkage.
- Centroid Method, menggunakan jarak euclid antar centroid.  $D(U, V) = d(yA, yB)$ ,
- Ward's Method = menggunakan kriteria ESS)

### Klasterisasi Fuzzy berbasis Prototipe (Prototype-Based Fuzzy Clustering)

Misalkan  $X = \{x_j | j=1 \dots N\}$  adalah sebuah subset terhingga dari sebuah ruang vektor berdimensi- $n$ .  $X$  adalah vector karakteristik. Misalkan  $C$  adalah banyaknya klaster, maka sebuah matriks  $U$  berukuran  $C \times N$  dimana  $U = [u_{ij}]$  disebut sebuah partisi  $C$  fuzzy terkendala dari  $X$  jika elemen dari  $U$  memenuhi

$$u_{ij} \in [0, 1] \text{ untuk semua } i, \quad 0 < \sum_{j=1}^N u_{ij} < N \text{ untuk semua } i, j \text{ dan } \sum_{j=1}^C u_{ij} = 1 \text{ untuk semua } j \quad (3)$$

$u_{ij}$  adalah derajat keanggotaan (*grade of membership*) dari  $x_j$  pada subset fuzzy ke  $i$  dari  $X$ . Algoritma fuzzy berbasis prototype mempartisi set data dengan meminimalisasi kriteria error kuadrat berikut :

$$J(B, U; X) = \sum_{i=1}^C \sum_{j=1}^N (u_{ij})^m d^2(x_j, \beta_i). \quad (4)$$

Pada persamaan (4),  $m \in (1, \infty)$  adalah sebuah eksponen pembobot,  $d^2(x_j, \beta_i)$  adalah jarak dari sebuah titik di  $x_j$  ke prototype  $\beta_i$ , dan  $B = (\beta_1, \dots, \beta_C)$  adalah sebuah tuple  $C$  dari prototype, yang masing-masing merupakan karakteristik dari setiap klaster dari sebanyak  $C$  klaster. Setiap prototype  $\beta_i$  berisi segugus parameter. Pada algoritma fuzzy C-means (FCM), yang merupakan salah satu algoritma yang sering digunakan dalam algoritma pengklasteran,  $\beta_i$  sederhananya adalah pusat klaster. Minimalisasi fungsi tujuan (*objective function*) pada persamaan (4) terhadap  $U$ , dengan fungsi kendala seperti pada (3) memberikan hasil:

$$\left. \begin{aligned} u_{ij} &= \frac{1}{\sum_{k=1}^C \left[ \frac{d^2(x_j, \beta_i)}{d^2(x_j, \beta_k)} \right]^{1/(m-1)}}, & \text{jika } I_j = \varnothing \\ u_{ij} &= 0 & i \notin I_j \\ \sum_{i \in I_j} u_{ij} &= 1, & i \in I_j \end{aligned} \right\}, & \text{jika } I_j \neq \varnothing \quad (5)$$

Dimana  $I = \{i | 1 \leq i \leq C, d^2(x_j, \beta_i) = 0\}$ .

Algoritma FCM biasanya menggunakan skema iterasi fixed point untuk menemukan solusi dalam permasalahan meminimalisasi fungsi tujuan.

### 3 METODE PENELITIAN

Metode penelitian menjelaskan rancangan kegiatan, ruang lingkup atau objek, bahan dan alat utama, tempat, teknik pengumpulan data, definisi operasional variable penelitian, dan teknik analisis. [Times New Roman, 11, normal].

Data yang digunakan pada makalah ini adalah data Susenas 2006, dengan perhatian utama pada variabel karakteristik perempuan. Variabel karakteristik perempuan tersebut digunakan sebagai dasar untuk mengelompokkan kabupaten-kabupaten di provinsi Jawa Timur. Untuk keperluan analisis klaster maka variabel karakteristik yang digunakan adalah:

- X1 = persentase jumlah perempuan yang telah mengenyam pendidikan dasar 9 tahun (minimal lulusan SMP)
- X2 = persentase jumlah perempuan yang berusia produktif (17 – 60 tahun)
- X3 = persentase jumlah perempuan yang berstatus menikah
- X4 = persentase jumlah perempuan yang bekerja dan mendapat upah

Pengklasteran dilakukan terhadap 23 kabupaten yang terdapat di Provinsi Jawa Timur. Data karakteristik perempuan pada 23 kabupaten di Jawa Timur disajikan pada Tabel 1.

Tabel 1 Karakteristik perempuan di 23 kabupaten pada Provinsi Jawa Timur berdasarkan data Susenas 2006

No	Kabupaten	X1	X2	X3	X4
1	Pacitan	26.15	59.11	53.14	4.16
2	Ponorogo	27.29	58.80	52.63	5.79
3	Trenggalek	27.22	61.13	53.46	6.09
4	Tulungagung	30.91	59.74	52.68	8.35
5	Blitar	31.27	60.69	54.65	7.15
6	Kediri	30.33	62.08	51.48	9.70
7	Malang	25.44	63.76	54.64	13.26
8	Lumajang	18.40	66.94	54.53	6.07
9	Jember	22.25	63.26	52.32	9.35
10	Banyuwangi	27.22	63.19	53.80	9.14
11	Bondowoso	10.22	63.57	55.70	4.02
12	Situbondo	20.97	64.73	56.07	7.11
13	Probolinggo	11.30	61.79	54.46	5.79
14	Pasuruan	17.82	63.41	49.49	10.41
15	Sidoarjo	49.69	65.87	50.08	15.06
16	Mojokerto	31.20	62.26	51.26	11.29
17	Jombang	32.03	61.45	50.57	7.79
18	Nganjuk	27.76	60.77	50.21	8.04
19	Madiun	27.94	61.59	49.05	6.43
20	Magetan	32.80	60.71	50.84	6.58
21	Ngawi	22.13	61.45	51.66	5.94
22	Bojonegoro	25.44	66.04	55.20	6.33
23	Tuban	20.84	64.13	51.68	6.50

### 4 HASIL DAN PEMBAHASAN

Analisis klaster berhirarkhi agglomerative selanjutnya diterapkan pada data di atas untuk tujuan pengklasteran kabupaten berdasarkan kemiripan karakteristik perempuan. Prosedur-prosedur yang

digunakan adalah single, average, complete, median, dan centroid linkage.

Analisis kluster agglomerative ini dilakukan sebagai langkah awal untuk membandingkan hasil klasifikasi berbagai prosedur untuk melihat karakteristik data. Untuk mempermudah perbandingan maka pada masing-masing prosedur ditetapkan 2 kluster sebagai hasilnya.

Hasil pengelompokan dengan teknik agglomerative tersebut disajikan pada Tabel 2.

Tabel 2 Status keanggotaan 23 kabupaten berdasarkan teknik agglomerative

Kabupaten/Kota	Single	Average	Centroid	Complete	Median
Pacitan	1	1	1	1	1
Ponorogo	1	1	1	1	1
Trenggalek	1	1	1	1	1
Tulungagung	1	1	1	1	1
Blitar	1	1	1	1	1
Kediri	1	1	1	1	1
Malang	1	1	1	1	1
Lumajang	1	1	1	1	1
Jember	1	1	1	1	1
Banyuwangi	1	1	1	1	1
Bondowoso	1	1	1	1	1
Situbondo	1	1	1	1	1
Probolinggo	1	1	1	1	1
Pasuruan	1	1	1	1	1
<b>Sidoarjo</b>	<b>2</b>	<b>2</b>	<b>2</b>	<b>2</b>	<b>2</b>
Mojokerto	1	1	1	1	1
Jombang	1	1	1	1	1
Nganjuk	1	1	1	1	1
Madiun	1	1	1	1	1
Magetan	1	1	1	1	1
Ngawi	1	1	1	1	1
Bojonegoro	1	1	1	1	1
Tuban	1	1	1	1	1

Baik dengan single, average, centroid, complete maupun median linkage, diperoleh 2 kluster yang sama. Kluster 1 berisi 22 kabupaten kecuali Sidoarjo yang ditandai dengan angka 1, sedangkan kluster 2 beranggotakan kabupaten Sidoarjo saja yang pada Tabel 2 di atas ditandai dengan angka 2.

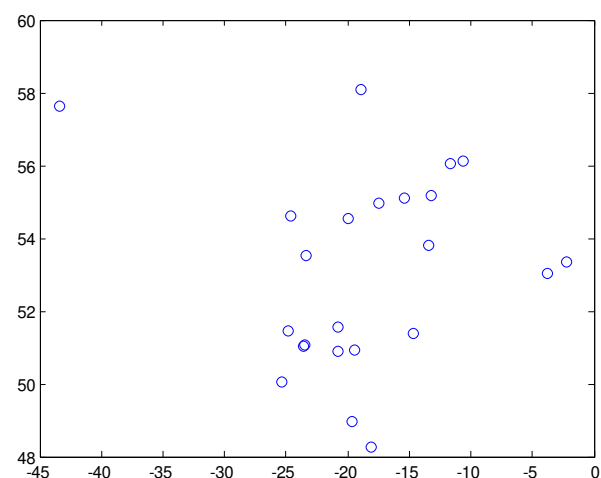
Analisis selanjutnya masih menggunakan metode *agglomerative* dengan single, average, centroid, complete dan median linkage tapi data pencilan dibuang, dalam hal ini data observasi kabupaten Sidoarjo (nomor 15) tidak digunakan lagi dalam analisis. Secara logis dapat diperkirakan bahwa akan terbentuk kluster baru yang berbeda dengan sebelumnya karena pada kluster sebelumnya (sebelum pencilan dibuang) terdapat satu kluster yang hanya terdiri dari satu anggota yaitu kabupaten Sidoarjo. Jika data kabupaten Sidoarjo dihapus otomatis akan terbentuk kluster baru yang benar-benar

berbeda. Hasil analisis kluster dengan single, average, centroid, complete, dan median linkage semuanya menunjukkan hasil yang sama. Dari 22 kabupaten yang diklusterkan, satu kluster beranggotakan 2 kabupaten yaitu kabupaten Bondowoso dan Probolinggo, sedangkan 20 kabupaten yang lain masuk pada kluster lainnya.

Struktur kluster yang diperoleh dengan sangat jelas menunjukkan bahwa terdapat masalah pencilan (*outlier*) pada data. Kenyataan ini membuat hasil pengklusteran sangat dipengaruhi oleh satu data yang memencil tersebut yaitu data karakteristik perempuan pada kabupaten Sidoarjo. Begitu data pencilan dihapus, maka terbentuk kluster baru yang berbeda dengan sebelumnya.

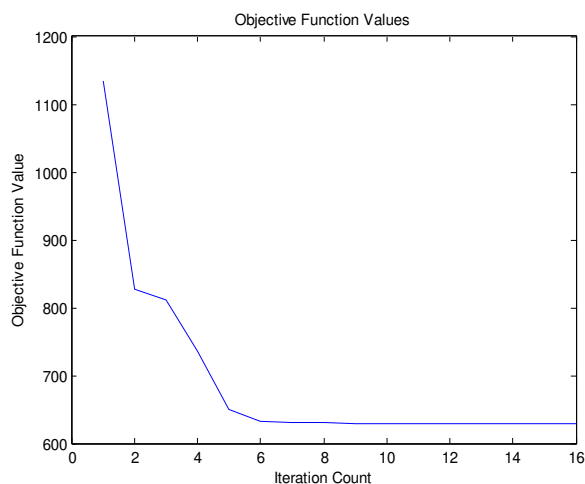
Hasil pengklusteran ini akan dibandingkan dengan hasil pengklusteran melalui teknik *prototype-based fuzzy clustering* dengan menggunakan algoritma FCM.

Pada tahap awal, data direduksi menjadi 2 variabel melalui analisis komponen utama. Dua skor pertama dari hasil analisis komponen utama selanjutnya digunakan sebagai dasar melakukan pengklusteran. Hasil plot terhadap data skor komponen utama disajikan pada Gambar 1.

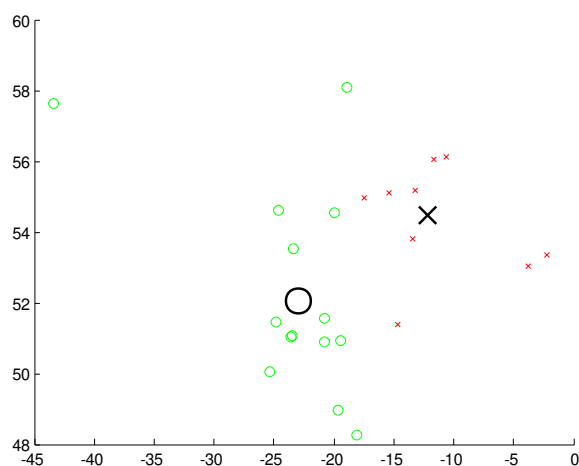


Gambar 1 Plot skor komponen utama

Dari Gambar 1 dapat dilihat bahwa terdapat 1 data yang memencil di sudut atas sebelah kiri. Selanjutnya pengklusteran melalui teknik *prototype-based fuzzy clustering* dengan menggunakan algoritma FCM diperoleh 2 kluster melalui iterasi sebanyak 16 kali. Iterasi berhenti ketika diperoleh nilai fungsi tujuan paling kecil. Berikut adalah gambaran nilai fungsi tujuan (*objective function*) dari iterasi 1 sampai iterasi ke 16 yang disajikan dalam Gambar 2 (a).



Gambar 2. Nilai-nilai fungsi tujuan hasil iterasi

Gambar 3. Plot hasil pengklasteran melalui *prototype-based fuzzy clustering* algoritma FCM

Hasil klasterisasi secara visual dapat disajikan pada Gambar 2(b). Klaster 1 beranggotakan kabupaten bernomor urut 1, 2, 3, 4, 5, 6, 7, 10, 15, 16, 17, 18, 19, dan 20. Sedangkan klaster 2 beranggotakan kabupaten-kabupaten bernomor urut 8, 9, 11, 12, 13, 14, 21, 22 dan 23. Dengan demikian kluster 1 terdiri dari kabupaten Pacitan, Ponorogo, Trenggalek, Tulungagung, Blitar, Kediri, Malang, Banyuwangi, Sidoarjo, Mojokerto, Jombang, Nganjuk, Madiun, dan Magetan. Sedangkan Klaster 2 terdiri dari kabupaten Lumajang, Jember, Bondowoso, Situbondo, Probolinggo, Pasuruan, Ngawi, Bojonegoro, dan Tuban.

Selanjutnya pengklasteran melalui teknik *prototype-based fuzzy clustering* dengan menggunakan algoritma FCM juga dilakukan dengan tidak melibatkan data pencilan. Jumlah kluster juga dibatasi 2 saja. Hasilnya adalah tidak ada perubahan yang signifikan mengenai kluster yang terbentuk. Klaster 1 terdiri dari kabupaten Pacitan, Ponorogo, Trengga-

lek, Tulungagung, Blitar, Kediri, Malang, Banyuwangi, Mojokerto, Jombang, Nganjuk, Madiun, Magetan dan Bojonegoro. Sedangkan Klaster 2 terdiri dari kabupaten Lumajang, Jember, Bondowoso, Situbondo, Probolinggo, Pasuruan, Ngawi, dan Tuban. Perbedaan dari kluster sebelumnya hanya pada kabupaten Bojonegoro yang pada pengklasteran sebelumnya masuk pada kluster 2.

Berdasarkan hasil analisis di atas maka diperoleh struktur kluster yang lebih *robust* jika menggunakan metode *prototype-based fuzzy clustering* algoritma FCM dibandingkan dengan metode *agglomerative*. Metode ini apat digunakan tanpa harus membuang data pencilan.

## 5 KESIMPULAN

Pengklasteran dengan metode *Agglomerative* sangat dipengaruhi oleh data pencilan. Keberadaan data pencilan tersebut membuat kluster-kluster yang dibentuk menjadi tidak *robust*. Begitu data pencilan dibuang, akan terbentuk kluster baru yang berbeda dengan sebelumnya.

Permasalahan mengenai keberadaan data pencilan tersebut dapat diatasi dengan metode *prototype-based fuzzy clustering*. Penggunaan metode ini membuat kluster-kluster yang terbentuk menjadi lebih *robust* dari pencilan.

## Rekomendasi

Untuk hasil yang lebih lengkap dapat dilakukan dengan metode lain seperti: Ohashi Algorithm (Ohashi, 1984), *noise clustering (NC) method*, *possibilistic clustering method* (Krishnapuram dan Keller, 1993), *the mountain method* (Yager dan Filev, 1994), *the deterministic annealing/least biased fuzzy clustering method* (Beni dan Liu, 1994), *Probabilistic Neural Network-based robust clustering* (Glotsos dkk, 2004).

## REFERENSI

- [1] Beni, G and X. Liu, 1994. "A least biased fuzzy clustering method," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 16, pp. 954-960, Sept. 1994.
- [2] Bezdek, J. C., *Pattern Recognition with Fuzzy Objective Function Algorithms*. New York: Plenum, 1981.
- [3] Chiu, S. and J. J. Cheng, "Automatic rule generation of fuzzy rule base for robot arm posture selection," in *Proc. NAFIPS Conf.*, San Antonio, TX, Dec. 1994, pp. 436-440.
- [4] Dillon, W.R. and Goldstein, M., 1984, *Multivariate Analysis and Application*. John Wiley and Son, New York

- [5] Glotsos, D., Jussi Tohka, Jori Soukka & Ulla Ruotsalainen. 2004. A New Approach to Robust Clustering by Density Estimation in an Autocorrelation Derived Feature Space. *Proceedings of the 6th Nordic Signal Processing Symposium - NORSIG 2004*, June 9 - 11, Espoo, Finland.
- [6] Härdle & Simar. 2007. *Applied Multivariate Statistical Analysis*, Second Edition. Springer, New York.
- [7] Johnson, R. & Wichern. 2007. *Applied Multivariate Statistical Analysis*, Sixth Edition. NJ.
- [8] Rencher, Alvin C., 2002. *Methods of Multivariate Analysis*, Second Edition. John Wiley & Sons, Inc., New York.
- [9] Krishnapuram, R. and J. M. Keller, "A possibilistic approach to clustering," *IEEE Trans. Fuzzy Syst.*, vol. 1, pp. 98–110, May 1993.
- [10] Ohashi, Y., 1984. "Fuzzy clustering and robust estimation," in *9th Meet. SAS Users Grp. Int.*, Hollywood Beach, FL, 1984.
- [11] Siswadi & Budi Suharjo. 1997. *Analisis Eksplorasi Peubah Ganda*. Jurusan Matematika FMIPA IPB, Bogor.
- [12] Supranto, J. 2010. *Analisis Multivariate*, arti dan interpretasi. Rineka Cipta, Jakarta.
- [13] Timm, Neil H., 2002. *Applied Multivariate Analysis*. Springer, New York.
- [14] Yager, R. R., and D. P. Filev, 1994. "Approximate clustering via the mountain method," *IEEE Trans. Syst. Man, Cybern.*, vol. 24, pp. 1279–1284, Aug. 1994. \_\_\_\_\_